

# Strange Weak Values

Akio Hosoya<sup>1,†</sup> and Yutaka Shikano<sup>1,2,\*</sup>

<sup>1</sup> Department of Physics, Tokyo Institute of Technology, Meguro, Tokyo 152-8551, Japan

<sup>2</sup> Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

E-mail: <sup>†</sup>ahosoya@th.phys.titech.ac.jp, <sup>\*</sup>shikano@mit.edu

**Abstract.** We develop a formal theory of the weak values with emphasis on the consistency conditions and a probabilistic interpretation in the counter-factual processes. We present the condition for the choice of the post-selected state to give a negative weak value of a given projection operator and strange values of an observable in general. The general framework is applied to Hardy's paradox and the spin 1/2 system to explicitly address the issues of counter-factuality and strange weak values. The counter-factual arguments which characterize the paradox specifies the pre-selected state and a complete set of the post-selected states clarifies how the strange weak values emerge.

PACS numbers: 03.65.Ta, 03.65.Ud, 03.65.Ca

Submitted to: *J. Phys. A: Math. Gen.*

## 1. Introduction

Many quantum paradoxes appear through attributes which are not actually measured. One may discard all such counter-factuality by insisting that attributes can have value only when they are actually measured. A well-known example is Hardy's paradox as we discuss later in detail. However, Aharonov and his colleagues have long claimed that there is a consistent and experimentally verifiable way to apply the so called *weak value* to counter-factual reasoning in quantum mechanics. The cost is that we have to allow strange values outside the range of the eigenvalue spectrum of the observable for the weak values. The concept of weak value was originally proposed by Aharonov and his collaborators [1, 2, 3, 4, 5] in terms of the weak measurement combined with pre-selected and post-selected states. Recently, experiments have verified aspects of weak value theory both optically [6, 7, 8, 9, 10, 11, 12, 13] and in the solid state qubits [14]. On the basis of the weak measurement, a formal theory of the weak value has also been developed [15, 16] ‡. However, we believe that the weak values have to be studied on its own right although its measurability is utmost important §.

The weak value of an observable  $A$  at time  $t$  is defined by

$${}_{\phi}\langle A \rangle_{\psi}^w := \frac{\langle \phi(t) | A | \psi(t) \rangle}{\langle \phi(t) | \psi(t) \rangle}, \quad \langle \phi(t) | \psi(t) \rangle \neq 0, \quad (1)$$

where the states  $|\psi(t)\rangle$  and  $\langle \phi(t)|$  are the solutions of the Schrödinger equation with given initial and final state conditions,  $|\psi(t_i)\rangle = |i\rangle$  and  $\langle \phi(t_f)| = \langle f|$ . This time-symmetric approach to quantum mechanics has been emphasized in Ref. [17]. Thus, the weak value shares the same spirit as the action principle in the sense that the both initial and final positions are fixed and can be written in terms of the Feynman path integral (also see Ref. [18]).

In this paper we develop a framework of weak value vectors where we fix a pre-selected state and consider a complete orthonormal set for post-selected states. We derive the two consistency conditions for the weak value vectors which enables us to interpret them as (complex) conditional probabilities for counter-factual processes in quantum mechanics. We shall explicitly give in what cases a weak value becomes strange. Two examples are for the demonstration; Hardy's paradox and the spin system. In the former the choice of the pre-selected state is unique from the setting of the paradox and the appearance of the “negative probability” is due to the counter-factuality. In the latter we discuss the case of non-commutative operators to weakly observe.

The organization of the present paper is as follows. In Sect. 2 we develop a formal aspect of weak values and give the two consistency conditions and a probabilistic interpretation of weak values. In Sect. 3, we point out that strange values result from negative weak values of projection operators, i.e., “negative probabilities”. We also show

‡ A compact review of the weak measurement can be seen in these references.

§ This point of view is strengthened by the recent analysis that demonstrates a reconstruction of weak values without using the weak measurements [37]. There is also a generalization of weak measurement using a weak entanglement [13].

the necessary and sufficient conditions for negative weak values of projection operators and give a formula for the most negative weak value. To demonstrate the analysis in Sect. 3, we examine Hardy's paradox in Sect. 4. as an arch-typical case and the spin 1/2 system in Sect. 5 to see the effect of non-commuting spin operators. Section 6 is devoted to summary.

## 2. Consistency and Probabilistic Interpretation of Weak Values

The weak values have been discussed mostly for commuting variables, e.g., optical path projection operators  $|m\rangle\langle m|$  ( $m = 1, 2, \dots, N$ ). Here, the consistency comes from the completeness relation,

$$\sum_m \phi \langle |m\rangle\langle m| \rangle_\psi^w = 1. \quad (\text{CONSISTENCY 1}) \quad (2)$$

That is, all the possible intermediate states add up to unity in the weak values. The consistency condition of this kind has been discussed in the literature, e.g., [15, 19]. According to the Born rule, which has been verified by numerous experiments, the expectation value of an observable  $A$  is given by

$$\begin{aligned} \langle \psi | A | \psi \rangle &= \sum_x \langle \psi | x \rangle \langle x | A | \psi \rangle \\ &= \sum_x |\langle \psi | x \rangle|^2 \frac{\langle x | A | \psi \rangle}{\langle x | \psi \rangle} \\ &= \sum_x |\langle \psi | x \rangle|^2 {}_x \langle A \rangle_\psi^w \\ &= \sum_x \text{Pr}(x) {}_x \langle A \rangle_\psi^w, \end{aligned} \quad (3)$$

provided that the set  $\{|x\rangle\}$  does not contain an element such that  $\langle x | \psi \rangle = 0$ . Here the  $\text{Pr}(x) = |\langle \psi | x \rangle|^2$  is the probability measure to find the state  $|x\rangle$  in the pre-selected state  $|\psi\rangle$ , which is defined with respect to a fixed complete set  $|x\rangle$  independent of the observable  $A$  itself.  ${}_x \langle A \rangle_\psi^w := \frac{\langle x | A | \psi \rangle}{\langle x | \psi \rangle}$  is the weak value of  $A$  with  $|\psi\rangle$  and  $\langle x|$  being the pre-selected and post-selected states, respectively. Comparing Eq. (3) with the expression for the expectation value  $\text{Ex}[A]$  of  $A$  in the standard probability theory,

$$\text{Ex}[A] = \sum_x \text{Pr}(x) h_A(x), \quad (4)$$

we see that the statistical average of the weak value becomes the expectation value in quantum mechanics and one may interpret the weak value as a complex random variable  $h_A(x)$  ||;

$$h_A(x) = {}_x \langle A \rangle_\psi^w \in \mathbb{C}. \quad (5)$$

We emphasize that the probability measure  $\text{Pr}(x) = |\langle \psi | x \rangle|^2$  does not depend on the observable  $A$ . Aharonov and Botero [20, Eq. (2.12)] derived the formula for the average

|| The Bayesian-like equation has been pointed out in Ref. [38].

of observables, which is close to our eq. (3), though the observable independence of the probability measure in our presentation seems new.

Further if the operator  $A$  is a projection operator  $A = |a\rangle\langle a|$ , then the identity (3) becomes an analog of the Bayesian rule in probability theory. The weak value  ${}_x\langle a|\langle a|\rangle_\psi^w$  can be interpreted as the conditional probability, which is in general complex, for the process: from the preselected state  $|\psi\rangle$  to the post selected state  $\langle x|$  via the intermediate state  $|a\rangle$ . The concept of negative probability is not new, e.g., [21, 22, 23]. The weak value defined by Eq. (1) is normally called the transition amplitude from the state  $|\psi\rangle$  to  $\langle\phi|$  via the intermediate state  $|a\rangle$  for  $A = |a\rangle\langle a|$ , the absolute value squared of which is the probability for the process. But the three references quoted above seem to suggest that they might be interpreted as probabilities in the case that the process is counter-factual, i.e., the case that the intermediate state  $|a\rangle$  is not projectively measured. The description of intermediate state  $|a\rangle$  in the present work is counter-factual or virtual in the sense that the intermediate state would not be observed by projective measurements. Feynman's example is the counter-factual "probability" for an electron to have its spin up in the  $x$ -direction and also spin down in the  $z$ -direction [22].

This is very different from the so-called Aharonov-Bergmann-Lebowitz (ABL) formula [24] for the conditional probability, which is the modulus squared of the weak value,

$$|{}_x\langle a|\langle a|\rangle_\psi^w|^2 = \frac{|\langle\phi|a\rangle|^2|\langle a|\psi\rangle|^2}{|\langle\phi|\psi\rangle|^2} = \frac{\Pr(a; \psi) \Pr(\phi; a)}{\Pr(\phi; \psi)}, \quad (6)$$

where  $\Pr(j; i) := |\langle j|i\rangle|^2$  is the transitional probability from the state  $|i\rangle$  to  $|j\rangle$ . The ABL formula has a probabilistic interpretation only if the intermediate state is projectively measured. For the counter-factual intermediate state the weak value plays a role of probability rather than amplitude from our view point. We will return to this counter-factuality of the weak values in the subsequent sections.

The stochastic theory tells us that the variance of the observable  $A$  is given by

$$\begin{aligned} Var(A) &= \sum_x |h_A(x)|^2 \Pr(x) - \left( \sum_x h_A(x) \Pr(x) \right)^2 \\ &= \sum_x |{}_x\langle A \rangle_\psi^w|^2 \Pr(x) - \left( \sum_x {}_x\langle A \rangle_\psi^w \Pr(x) \right)^2 \\ &= \langle \psi | (A - \langle \psi | A | \psi \rangle)^2 | \psi \rangle. \end{aligned} \quad (7)$$

The last line comes from a straightforward calculation. The result coincides with the conventional "standard deviation squared" in quantum mechanics, which we claim is derived from the identification  $h_A(x) = {}_x\langle A \rangle_\psi^w$ .

Equation (3) is further generalized to

$$\sum_x ({}_x\langle A \rangle_\psi^w)^* |{}_x\langle \psi \rangle|^2 {}_x\langle B \rangle_\psi^w = \langle \psi | AB | \psi \rangle. \quad (\text{CONSISTENCY 2}) \quad (8)$$

The proof is similar to that given for Eq. (3). The formula (8) reduces to Eq. (3) for  $B = 1$  and for  $A = B$  to the  $\langle \psi | A^2 | \psi \rangle$  term of Eq. (7). In particular, if  $A$  and  $B$  are

orthogonal projection operators, the right hand side of Eq. (8) becomes  $\delta_{AB}\langle\psi|A|\psi\rangle$ , which gives an orthonormality condition for the weak values .

Let  $A = |a\rangle\langle a|$  and  $B$  be two in general non-commuting projection operators. If we adopt the previous interpretation of weak values as a conditional complex probability, the left hand side of Eq. (8) implies that the right hand side  $\langle\psi|AB|\psi\rangle$  has a meaning of the joint probability for  $B$  and then  $A$  to occur if it is real, i.e.,  $A$  and  $B$  are commutable in the sense of expectation value ¶. Putting the other way, we confirm this by the identity,

$$\langle\psi|AB|\psi\rangle = |\langle a|\psi\rangle|^2 \frac{\langle a|B|\psi\rangle}{\langle a|\psi\rangle} = |\langle a|\psi\rangle|^2 {}_a\langle B\rangle_\psi^w. \quad (9)$$

Steinberg [25, 26] addressed the issue of the tunneling traversal time from the view point of the weak measurement in which he proposed the joint and conditional probabilities for the counter-factual cases in quantum mechanics. We regard the relation (9) of the joint and conditional probabilities as consistency conditions for a set of weak values in a framework where the weak value is promoted to a fundamental concept independent of the weak measurement model.

It is interesting to point out that the quantity  $\langle\psi|(A-B)^2|\psi\rangle$ , which is an indicator of the difference of the two quantities  $A$  and  $B$  in the state  $|\psi\rangle$ , can be evaluated by the experimentally accessible weak values via the following formula:

$$\sum_x ({}_x\langle A\rangle_\psi^w - {}_x\langle B\rangle_\psi^w)^* |\langle x|\psi\rangle|^2 ({}_x\langle A\rangle_\psi^w - {}_x\langle B\rangle_\psi^w) = \langle\psi|(A-B)^2|\psi\rangle. \quad (10)$$

With  $|\psi\rangle$  being the pre-selected state, the probability to get a post-selected state  $|x\rangle$  is  $\text{Pr}(x) = |\langle x|\psi\rangle|^2$  so that the left hand side can be expressed as

$$\sum_x \text{Pr}(x) |h_A(x) - h_B(x)|^2. \quad (11)$$

which can be computed exploiting all the data exhausting all possible post-selected states. In terms of the weak measurements, a particular post-selection  $x$  shifts the center of the distribution of the pointer position by  $h_A(x)$  when  $A$  is measured and by  $h_B(x)$  for  $B$ . The difference  $h_A(x) - h_B(x)$  varies depending on the post-selected state  $|x\rangle$ . The weighted sum of the square of the difference over the post-selected states gives the quantity  $\langle\psi|(A-B)^2|\psi\rangle$ . Note that the probability measure defined independently of the observed quantities helps to give a symmetric expression for the two observables  $A$  and  $B$ . We would like to remark that the following simple fact.

**Fact 1.** *Fix the quantum state  $|\psi\rangle$ . The observables  $A$  and  $B$  are equivalent for the quantum state  $|\psi\rangle$  if and only if*

$$\langle\psi|(A-B)^2|\psi\rangle = 0. \quad (12)$$

¶ The authors thank Professor Ozawa for this comment who has formulated the state-dependent joint probability [39].

Note that this condition implies that

$${}_{\phi}\langle(A-B)\rangle_{\psi}^w = 0 \quad (13)$$

for any post-selected states  $\langle\phi|$ . That is, two observables are equivalent if their weak value vectors are identical. This fact will be used to check whether the state  $|\psi\rangle$  is the correct one for an experimental setup in Sect. 4.1.

A keen reader may notice that only the real part of the weak value is relevant in the Bayesian-like rule (3) so that there is an option that we interpret only the real part as a conditional probability. Then it would be interesting to consider the restricted cases where the weak values are real. This also demands the reality of the joint probability (9), which implies the commutativity of the observables  $A$  and  $B$  relative to the pre-selected state  $|\psi\rangle$ , i.e., there is no Heisenberg-Robertson uncertainty  $\Delta A \Delta B \geq 0$  for the state  $|\psi\rangle$ , where  $\Delta A$  and  $\Delta B$  are the square root of the variances of  $A$  and  $B$ . A simple example is the case:  $A = \sigma_x, B = \sigma_y$  and  $|\psi\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$  which gives  $\Delta\sigma_x\Delta\sigma_y = |\langle\psi|\sigma_z|\psi\rangle| = 0$ . Namely, there is no uncertainty bound for the  $\sigma_x$  and the  $\sigma_y$  because they are commuting in the state  $|\psi\rangle$ . Hereafter we will be mainly concerned with the real weak value unless otherwise stated.

### 3. Condition for a Strange Weak Value

When some of the weak values of the projection operators exceed unity, at least one of the other weak values, say  $\text{Pr}_m := {}_f\langle|m\rangle\langle m|_i\rangle^w$  becomes negative. One might interpret the  $\text{Pr}_m$  as a “negative” conditional probability for the path  $|i\rangle \rightarrow |m\rangle \rightarrow |f\rangle$ . We can see the negative conditional probability is also responsible for the various “strange” weak values, e.g., the weak value of the  $z$ -component of the spin of a Dirac particle can be very large, say  $100\hbar [1]^+$ . Let the spectral decomposition of an observable  $A$  be

$$A = \sum_{n=1}^N a_n |n\rangle\langle n|, \quad (14)$$

where the eigenvalues  $a_n$  obey the order relation  $a_N > a_{N-1} > \dots > a_1$  and  $|n\rangle$  is the eigenstate with the eigenvalue  $a_n$ . We concentrate on the maximum and minimum states,  $|N\rangle$  and  $|1\rangle$ . The pre-selected state is assumed to be

$$|\psi\rangle = \alpha|N\rangle + \beta|1\rangle, \quad \alpha, \beta \in \mathbb{R} \quad (15)$$

Choose the post-selected states  $|\phi\rangle$  as

$$|\phi\rangle = \frac{|N\rangle + |1\rangle}{\sqrt{2}} \quad (16)$$

to obtain the weak value of  $A$  as

$${}_{\phi}\langle A \rangle_{\psi}^w = a_N {}_{\phi}\langle |N\rangle\langle N| \rangle_{\psi}^w + a_1 {}_{\phi}\langle |1\rangle\langle 1| \rangle_{\psi}^w = a_N + (a_1 - a_N) {}_{\phi}\langle |1\rangle\langle 1| \rangle_{\psi}^w. \quad (17)$$

<sup>+</sup> Of course, this is not mathematically strange at all. It is strange only if one insists on the probabilistic interpretation of the weak values.

This can be larger than the maximum eigenvalue  $a_N$  of  $A$  for negative  ${}_\phi\langle 1|1\rangle_\psi^w$ , i.e., for  $\alpha/\beta < -1$ . Similarly, we can analyze the case that the weak value is smaller than the minimum eigenvalue. The point is that at least one of the weak values of the projection operators is negative to have a weak value outside the range of the spectrum. The converse also holds, i.e., if the weak values of all the projection operators are positive, the weak value of the observable  $A$  is not strange.

In the weak measurement the post-selection biases the original probability distribution to push the mean value of the pointer position sometimes to a strange value outside the spectrum. There should be the optimal post-selection which shifts the pointer position most. We are going to find the condition for a weak value to be strange. From the argument stated above, we see that it is sufficient to consider one of the eigenvectors of the observable  $|n\rangle$  and find the condition for the weak value of the projection operator  $|n\rangle\langle n|$  to be negative.

Let us consider the weak value of a projection operator  $|n\rangle\langle n|$ ,

$$w(\phi) := {}_\phi\langle n|n\rangle_\psi^w = \frac{\langle\phi|n\rangle\langle n|\psi\rangle}{\langle\phi|\psi\rangle}. \quad (18)$$

We are going to find the optimal post-selection to obtain the most negative weak value  $w(\phi)$  for a given pre-selected state and the projection operator  $|n\rangle\langle n|$  to be weakly measured. Note that the probability  $\text{Pr}(\phi) := |\langle\phi|\psi\rangle|^2$  to obtain  $|\phi\rangle$  has to be finite for the weak value to be finite. Otherwise, the weak measurement would produce null results. The smaller the probability  $\text{Pr}(\phi)$  is, the more we need to repeat the weak measurements. We fix the value  $\text{Pr}(\phi) = |\langle\phi|\psi\rangle|^2 = \cos^2 \xi$  with  $\xi \in (0, \pi/2)$  \*.

To simplify the argument we also assume that all the inner products are real so that we can visualize the state vectors in the multi-dimensional real vector space. In particular, the post-selected state  $|\phi\rangle$  is lying on the cone defined by  $\langle\phi|\psi\rangle = \cos \xi$ . Let the angle between  $|n\rangle$  and  $|\psi\rangle$  be  $\theta_n \in (0, \pi/2)$ . Then the most negative weak value  $w(\phi)$  is given by a state  $|\phi\rangle$  which is the intersection of the cone  $\langle\phi|\psi\rangle = \cos \xi$  and the plane defined by the two vectors  $|\psi\rangle$  and  $|n\rangle$ . Then the weak value  $w(\phi)$  becomes

$$w(\phi) = \frac{\langle\phi|n\rangle\langle n|\psi\rangle}{\langle\phi|\psi\rangle} = \frac{\cos(\theta_n + \xi) \cos \theta_n}{\cos \xi}. \quad (19)$$

This can be negative if the angle between  $|n\rangle$  and  $|\phi\rangle$ ,  $\theta_n + \xi$  is obtuse, i.e.,

$$\xi + \theta_n > \pi/2, \quad (20)$$

provided that the angles  $\xi, \theta_n$  are acute, which can be assumed without loss of generality. Note that the above condition cannot be met if the angle  $\xi$  is too small, i.e., the probability  $\text{Pr}(\phi) = \cos^2 \xi$  is too close to unity. There is a trade-off between the labor and the fruit, i.e., the repetition of the weak measurements and the large negative weak value.

\* However, there is a proposal of rather sophisticated single-shot weak measurement, which is out of scope of the present work [40].

We turn to an interesting but more technical problem to find the most strange weak value of the observable  $A$ ,

$${}_{\phi}\langle A \rangle_{\psi}^w = \frac{\sum_{n=1}^N a_n \langle \phi|n \rangle \langle n|\psi \rangle}{\langle \phi|\psi \rangle}, \quad (21)$$

under the constraints:  $\langle \phi|\psi \rangle = \cos \xi$ ,  $\langle \phi|\phi \rangle = 1$ . Introducing the corresponding Lagrange multipliers  $\lambda$  and  $\mu$ , and following the standard variation method, we obtain the equation for  $\langle n|\phi \rangle$  as

$$\sum_{n=1}^N a_n |n\rangle \frac{\langle n|\psi \rangle}{\cos \xi} - \lambda |\psi \rangle - \mu |\phi \rangle = 0. \quad (22)$$

The formal solution is

$$|\phi \rangle = \frac{1}{\mu} \left[ \sum_{n=1}^N a_n |n\rangle \frac{\langle n|\psi \rangle}{\cos \xi} - \lambda |\psi \rangle \right], \quad (23)$$

where the multipliers  $\lambda$  and  $\mu$  are given by

$$\lambda = \sum_{n=1}^N \frac{a_n}{\sin^2 \xi} \left[ \frac{\langle n|\psi \rangle^2}{\cos \xi} - \langle \phi|n \rangle \langle n|\psi \rangle \right], \quad (24)$$

$$\mu = - \sum_{n=1}^N \frac{a_n}{\sin^2 \xi} \left[ \langle n|\psi \rangle^2 - \frac{\langle \phi|n \rangle \langle n|\psi \rangle}{\cos \xi} \right]. \quad (25)$$

The post-selection by the state (23) gives the most strange weak value. By projecting out this state by the base  $|n\rangle$  we have a closed set of algebraic equations for  $\langle \phi|n \rangle$ :

$$\begin{aligned} & a_n \langle \phi|n \rangle \sin^2 \xi \\ &= \langle \phi|n \rangle \left[ \sum_{m=1}^N a_m (\langle \phi|m \rangle - \langle \psi|m \rangle \cos \xi) \langle \phi|m \rangle \right] \\ & \quad - \langle \psi|n \rangle \left[ \sum_{m=1}^N a_m (\langle \phi|m \rangle \cos \xi - \langle \psi|m \rangle) \langle \phi|m \rangle \right] \quad (n = 1 \dots N). \end{aligned} \quad (26)$$

For the particular case of single state  $|n\rangle$ , the previous result is reproduced by one of the solutions for  $\langle \phi|n \rangle = \cos(\theta + \xi)$ .

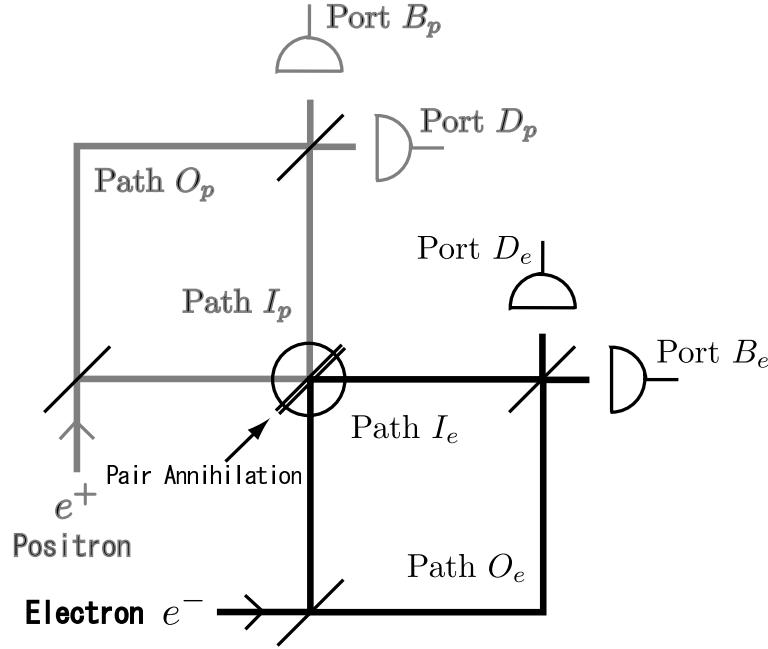
We are going to demonstrate the above statement in two explicit examples; Hardy's setting and the spin 1/2 system. In the first example, the weak values are real and the "negative probability" appears, while in the second spin system complex weak values appear and the spin  $x$ -component becomes large, say,  $100\hbar$ .

#### 4. Hardy's Paradox

Let us recall what Hardy's paradox is.

**Hardy's Paradox [27].** *The two Mach-Zehnder interferometers for an electron and a positron are combined in such a way that one of the arms, say,  $I_e$  of the interferometer for the electron is crossed with one of the arms  $I_p$  for the positron. If the electron and*





**Figure 1.** Hardy's setup: The Hardy paradox is explained in the main text.

the positron meet at the crossing they are supposed to annihilate each other. The setting is illustrated in Fig. 1.

- 1) Both the electron and positron cannot be in the inner path  $I_e$  and  $I_p$  at the same time since they would annihilate each other. Therefore, either electron or positron is in the outer path  $O_e$  or  $O_p$ .
- 2) Let the positron be in the path  $O_p$  so that the electron is not affected by the positron. The Mach-Zehnder interferometer works as usual; the intermediate state is  $O_p(I_e + O_e)$  (a shorthand for  $|O_p\rangle|\frac{(I_e+O_e)}{\sqrt{2}}\rangle$  which we use in this section for notational simplicity) and therefore the electron registers a click at detector  $B_e$ , while the positron clicks either the port  $B_p$  or  $D_p$ . Similarly, for the electron in  $O_e$ , the intermediate state will be  $(I_p + O_p)O_e$  and therefore the positron registers a click at detector  $B_p$ , while the electron clicks either  $B_e$  or  $D_e$ . In either case, only possible cases are:  $B_pB_e, D_pB_e, B_pD_e$ , but  $D_pD_e$  is not.
- 3) However, quantum mechanics tells us that  $D_pD_e$  clicks with the probability  $1/12$ , which has also been verified by experiments.

The line of reasoning 2) is not a consequence of measurements but an inference by the counter-factual argument [28]. For example, if the positron is found in  $O_p$  by measurement, the electron is definitely in  $(I_e + O_e)$  ending up with  $B_e$ , which can be demonstrated as an actual experiment. Another different experiment would be the one starting from the fact that the electron found in  $O_e$ . But one cannot infer anything about the non-measured quantity by combining the two different experiments of non-commuting observables  $O_p(I_e + O_e)$  and  $(I_p + O_p)O_e$ . The most obvious way to resolve

Hardy's paradox is to dismiss the counter-factual reasoning all together.

However, Aharonov and his colleagues have long advocated the use of weak values which are experimentally accessible to provide experimental credence to the counter-factual arguments [29]. The weak value gives a consistent description of the quantum evolution of the intermediate states at the cost of strange values, i.e., the value outside the range of the spectrum of observable. Actually they have been recently demonstrated in the optical settings [12, 13]. It should be noted that the Hardy paradox has also been analyzed in the different contexts [30, 31, 32, 33, 34, 35]. We are going to closely examine the counter-factual step 2) by looking at the corresponding weak values. We will see what is missing in the argument of 2) and how the counter-factual argument works in terms of weak values.

#### 4.1. Pre-selected State of the Hardy Paradox

In this section, we justify the standard initial quantum state,

$$|\psi\rangle = \frac{1}{\sqrt{3}}(|I_p O_e\rangle + |O_p I_e\rangle + |O_p O_e\rangle), \quad (27)$$

on the basis of the counter-factual arguments 2) in the original Hardy paradox and then show how to experimentally check whether the initial quantum state is the right hand side of Eq. (27) or not by evaluating the weak values.

The pre-selected state is in general given by

$$|\psi\rangle = \eta|I_p I_e\rangle + x|I_p O_e\rangle + y|O_p I_e\rangle + z|O_p O_e\rangle \quad (28)$$

with the coefficients  $\eta, x, y, z \in \mathbb{C}$  satisfying the normalization condition,  $|\eta|^2 + |x|^2 + |y|^2 + |z|^2 = 1$ . We are going to determine the coefficients by the following counter-factual arguments. Since the electron going through the path  $O_e$  does not affect the path of the positron, the projection operators  $P[O_p(I_e + O_e)]$  and  $P[O_p \otimes id.]$  are equivalent in the initial quantum state  $|\psi\rangle$ . Here  $id.$  denotes the identity operator. Similarly,  $P[(I_p + O_p)O_e]$  and  $P[id. \otimes O_e]$  are equivalent in  $|\psi\rangle$ . Furthermore, since the electron and the positron simultaneously going through the paths  $I_p$  and  $I_e$  would annihilate each other,  $P[I_p O_e]$  and  $P[I_p \otimes id.]$  and  $P[O_p I_e]$  and  $P[id. \otimes I_e]$  are also equivalent for  $|\psi\rangle$ . According to Fact 1, the above counter-factual arguments demand

$$\begin{aligned} & \langle\psi|(P[O_p(I_e + O_e)] - P[O_p \otimes id.])^2|\psi\rangle \\ &= \langle\psi|(P[O_p(I_e + O_e)] - P[O_p \otimes id.])|\psi\rangle = 0 \end{aligned} \quad (29)$$

$$\begin{aligned} & \langle\psi|(P[(I_p + O_p)O_e] - P[id. \otimes O_e])^2|\psi\rangle \\ &= \langle\psi|(P[(I_p + O_p)O_e] - P[id. \otimes O_e])|\psi\rangle = 0 \end{aligned} \quad (30)$$

$$\begin{aligned} & \langle\psi|(P[I_p O_e] - P[I_p \otimes id.])^2|\psi\rangle \\ &= \langle\psi|(P[I_p O_e] - P[I_p \otimes id.])|\psi\rangle = 0 \end{aligned} \quad (31)$$

$$\begin{aligned} & \langle\psi|(P[O_p I_e] - P[id. \otimes I_e])^2|\psi\rangle \\ &= \langle\psi|(P[O_p I_e] - P[id. \otimes I_e])|\psi\rangle = 0 \end{aligned} \quad (32)$$

From Eqs. (31) or (32), we obtain  $\eta = 0$ . From Eqs. (29) and (30), we obtain  $x = y = z$ . Taking into account the normalization condition, we obtain the standard Hardy initial state (27). The second equations of Eqs. (29), (30), (31) and (32) also imply that the initial quantum state can be experimentally tested by the joint expectation value of the paths for the electron and the positron.

Furthermore, we will show that this initial state can also be verified by the joint weak measurement, that is, the state can be constructed from a set of weak values. The inner product of the pre- and post-selected states are

$$\zeta_{D_p D_e} := \langle D_p D_e | \psi \rangle = \frac{1}{2}(\eta - x - y + z), \quad (33)$$

$$\zeta_{D_p B_e} := \langle D_p B_e | \psi \rangle = \frac{1}{2}(\eta + x - y - z), \quad (34)$$

$$\zeta_{B_p D_e} := \langle B_p D_e | \psi \rangle = \frac{1}{2}(\eta - x + y - z), \quad (35)$$

$$\zeta_{B_p B_e} := \langle B_p B_e | \psi \rangle = \frac{1}{2}(\eta + x + y + z). \quad (36)$$

We obtain the following table of weak values for the pre-selected state  $|\psi\rangle$  and the post-selected state  $|\phi\rangle$ :

Weight $\phi$	$\zeta_{D_p D_e}^2$ $D_p D_e$	$\zeta_{D_p B_e}^2$ $D_p B_e$	$\zeta_{B_p D_e}^2$ $B_p D_e$	$\zeta_{B_p B_e}^2$ $B_p B_e$
$O_p(I_e + O_e)$	0	$-\frac{y+z}{\zeta_{D_p B_e}}$	0	$\frac{y+z}{\zeta_{B_p B_e}}$
$(I_p + O_p)O_e$	0	0	$-\frac{x+z}{\zeta_{B_p D_e}}$	$\frac{x+z}{\zeta_{B_p B_e}}$
$I_p O_e$	$-\frac{x}{2\zeta_{D_p D_e}}$	$\frac{x}{2\zeta_{D_p B_e}}$	$-\frac{x}{2\zeta_{B_p D_e}}$	$\frac{x}{2\zeta_{B_p B_e}}$
$O_p I_e$	$-\frac{y}{2\zeta_{D_p D_e}}$	$-\frac{y}{2\zeta_{D_p B_e}}$	$\frac{y}{2\zeta_{B_p D_e}}$	$\frac{y}{2\zeta_{B_p B_e}}$
$O_p \otimes id.$	$\frac{-y+z}{2\zeta_{D_p D_e}}$	$\frac{-y-z}{2\zeta_{D_p B_e}}$	$\frac{y-z}{2\zeta_{B_p D_e}}$	$\frac{y+z}{2\zeta_{B_p B_e}}$
$I_p \otimes id.$	$\frac{\eta-x}{2\zeta_{D_p D_e}}$	$\frac{\eta+x}{2\zeta_{D_p B_e}}$	$\frac{\eta-x}{2\zeta_{B_p D_e}}$	$\frac{\eta+x}{2\zeta_{B_p B_e}}$
$id. \otimes O_e$	$\frac{-x+z}{2\zeta_{D_p D_e}}$	$\frac{x-z}{2\zeta_{D_p B_e}}$	$\frac{-x-z}{2\zeta_{B_p D_e}}$	$\frac{x+z}{2\zeta_{B_p B_e}}$
$id. \otimes I_e$	$\frac{\eta-y}{2\zeta_{D_p D_e}}$	$\frac{\eta-y}{2\zeta_{D_p B_e}}$	$\frac{\eta+y}{2\zeta_{B_p D_e}}$	$\frac{\eta+y}{2\zeta_{B_p B_e}}$

where the “Weight” is the probability of transition from the pre-selected state  $|\psi\rangle$  to the post-selected state  $|\phi\rangle$ . Here, the column of the post-selected state  $|\phi\rangle$  shows the weak values for each observable, e.g.  $O_p(I_e + O_e)$ . Interestingly our table (37) is similar to the one in Feynman’s paper [22]. From the same counter-factual arguments as before together with Eq. (13), we see that

$$InPositron(D_p D_e) := {}_{D_p D_e} \langle I_p O_e \rangle_\psi^w - {}_{D_p D_e} \langle I_p \otimes id. \rangle_\psi^w = 0, \quad (38)$$

$$InEletron(D_p D_e) := {}_{D_p D_e} \langle O_p I_e \rangle_\psi^w - {}_{D_p D_e} \langle id. \otimes I_e \rangle_\psi^w = 0, \quad (39)$$

$$OutPositron(D_p D_e) := {}_{D_p D_e} \langle O_p(I_e + O_e) \rangle_\psi^w - {}_{D_p D_e} \langle O_p \otimes id. \rangle_\psi^w = 0, \quad (40)$$

$$OutEletron(D_p D_e) := {}_{D_p D_e} \langle (I_p + O_p)O_e \rangle_\psi^w - {}_{D_p D_e} \langle id. \otimes O_e \rangle_\psi^w = 0. \quad (41)$$

Using the results of the above table in Eqs. (38) and (39) yields

$$\eta - x = -x \quad \text{or} \quad \eta - y = -y \quad (42)$$

to obtain that the coefficient of  $|I_p I_e\rangle$  must be  $\eta = 0$ . Similarly from Eqs. (40) and (41), we have

$$-y + z = 0 \quad \text{and} \quad -x + z = 0 \quad (43)$$

to obtain the coefficients  $x = y = z = 1/\sqrt{3}$  by the normalization condition.

We therefore conclude that

**Fact 2.** *The following conditions*

$$(i) \text{ InPositron}(D_p D_e) = \text{InElectron}(D_p D_e) = 0,$$

$$(ii) \text{ OutPositron}(D_p D_e) = \text{OutElectron}(D_p D_e) = 0,$$

*are satisfied if and only if the pre-selected state is Eq. (27).*

Note that we only need a particular post-selected state  $|D_p D_e\rangle$  to verify whether a prepared setup gives the standard Hardy's initial state (27) or not.

It is interesting to point out that the relations between the weak values can be used to evaluate the weak values of the non-local operators like  $I_p O_e, O_p I_e, O_p(I_e + O_e), (I_p + O_p)O_e$  using those of the local operators  $I_p \otimes id., id. \otimes I_e, O_p \otimes id., id. \otimes O_e$ , respectively, provided that the pre-selected state is Eq. (27).

#### 4.2. Weak values in the Hardy Paradox

For a given pre-selected state  $|\psi\rangle$  and a post-selected state  $\langle\phi|$  the weak value of an observable  $A$  is defined by

$${}_{\phi}\langle A \rangle_{\psi}^w = \frac{\langle\phi|A|\psi\rangle}{\langle\phi|\psi\rangle}. \quad (44)$$

This can be considered as a vector with component index  $\phi$  while the preselected state  $|\psi\rangle$  is fixed. This vector notation is convenient to find the condition for the strange weak values as we will see below. The pre-selected state in Hardy's paradox is

$$|\psi\rangle = \frac{|I_p O_e + O_p I_e + O_p O_e\rangle}{\sqrt{3}}. \quad (45)$$

as we have discussed in the previous subsection. The possible post-selected states are

$$|D_p D_e\rangle, |D_p B_e\rangle, |B_p D_e\rangle, |B_p B_e\rangle. \quad (46)$$

If we choose a set of non-commuting projection operators:

$$P[O_p(O_e + I_e)], P[(O_p + I_p)O_e], P[O_p I_e + I_p O_e], \quad (47)$$

then, the weak values for them are real and summarized in following table:

Weight $\phi$	1/12 $D_p D_e$	1/12 $D_p B_e$	1/12 $B_p D_e$	3/4 $B_p B_e$	Average
$O_p(I_e + O_e)$	0	2	0	2/3	2/3
$(I_p + O_p)O_e$	0	0	2	2/3	2/3
$I_p O_e + O_p I_e$	2	0	0	2/3	2/3

(48)

where “Average” is defined as the sum of the weak value times “Weight” over the post-selected states. Note that all the components of the weak values of the non-commuting projection operators  $P[I_p O_e + O_p I_e]$ ,  $P[O_p(I_e + O_e)]$ ,  $P[(I_p + O_p)O_e]$  are positive. The first two cases,  $O_p(I_e + O_e)$  and  $(I_p + O_p)O_e$  have played important roles in the previous counter-factual arguments. The non-commutativity is explicit as the non-orthogonality of the vectors, e.g.,  $(1, 0, 0, 1/3)$  and  $(0, 1, 0, 1/3)$ . The weight is defined by the square of the inner product as  $1/12, 1/12, 1/12$  and  $3/4$ . Furthermore, the average means the weighted average of the weak value for each row. It is interesting to point out that if  $D_e$  and  $D_p$  click the intermediate state should be definitely an entangled state  $I_p O_e + O_p I_e$ . If one tries to include the path  $O_p I_e + I_p O_e$ , the argument would become more counter-factual, because any local measurement would destroy this entangled state. This nonlocal state is one of the origin of Hardy’s paradox parallel to the Einstein-Podolsky-Rosen (EPR) paradox [36], because the classical argument 2) excludes this possibility of entanglement. These arguments are counter-factual in the sense that the identity of the intermediate states cannot be revealed by projective measurements. However, the weak measurement can do the job.

One might be curious about the strange value 2, e.g., for the projection operator  $O_p(I_e + O_e)$  and the post-selected state  $B_p D_e$  in Table (48). This can be explained in a counter-factual way by looking at the projection operators orthogonal to  $O_p(I_e + O_e)$ . For example, take  $I_p O_e$  orthogonal to  $O_p(I_e + O_e)$ . The table becomes

Weight	1/12	1/12	1/12	3/4	
$\phi$	$D_p D_e$	$D_p B_e$	$B_p D_e$	$B_p B_e$	Average
$O_p(I_e + O_e)$	0	2	0	2/3	2/3
$I_p O_e$	1	-1	1	1/3	1/3
Sum	1	1	1	1	1

(49)

We immediately see that the weak value  $-1$  for  $I_p O_e$  with  $D_p B_e$  partially cancels 2 for  $O_p(I_e + O_e)$  to give 1 as the sum. Note that the other orthogonal projection operator to  $O_p(I_e + O_e)$  gives the zero weak values supporting the counter factual argument 2) of Hardy’s paradox. Obviously, we can make the completely parallel argument for  $(I_p + O_p)O_e$  with the post-selected state  $B_p D_e$ .

It is interesting to look at the factual cases when we measure the intermediate states corresponding to the weak values in Table (49). Suppose we locally measure the intermediate state of the positron to find  $O_p$  implying the electron state is  $|I_e + O_e\rangle/\sqrt{2}$ , the probability of which is  $2/3$  according to the Born rule  $|\langle \frac{O_p(I_e + O_e)}{\sqrt{2}} | \psi \rangle|^2$ . Either positron detector  $B_p$  or  $D_p$  clicks with probability  $1/2$  and only  $B_e$  of the electron detectors clicks. Thus, the probabilities are 0 for  $D_p D_e$  and  $B_p D_e$ , while  $(1/2) \times (2/3) = 1/3$  for  $D_p B_e$  and  $B_p B_e$ . Similarly, in the case when  $I_p$  is the positron intermediate state, the probabilities are  $(1/4) \times (1/3) = 1/12$  for all  $D_p D_e, D_p B_e, B_p D_e$  and  $B_p B_e$ . These factual results can be also reproduced by the ABL formula [24], i.e., the absolute value squared of the weak values multiplied by the weight of the post-selected state. In the case of  $O_p$  for the positron, the ABL formula gives  $0, 2^2 \times (1/12) = 1/3, 0$

and  $(2/3)^2 \times (3/4) = 1/3$  for  $D_p D_e, D_p B_e, B_p D_e$  and  $B_p B_e$ , respectively. In the case of  $I_p$ , it reads  $1^2 \times (1/12) = 1/12, (-1)^2 \times (1/12) = 1/12, 1^2 \times (1/12) = 1/12$  and  $(1/3)^2 \times (3/4) = 1/12$  for  $D_p D_e, D_p B_e, B_p D_e$  and  $B_p B_e$ , respectively. Note that this coincidence is possible only for the factual and mutually orthogonal set of projection operators and does not work for the counter-factual cases. In the latter case we have to resort to fully quantum mechanical computations rather than the semi-classical counter-factual arguments. For the factual cases, that is, when we projectively measure the intermediate state  $|n\rangle$ , we can confirm the ABL formula for the partial probabilities by Eq. (6).

So far we have fixed the post-selected states as  $D_p D_e, D_p B_e, B_p D_e$  and  $B_p B_e$ . Here we consider a different choice of post-selected state  $|\phi\rangle$  keeping the same pre-selected state  $|\psi\rangle = (|I_p O_e\rangle + |O_p I_e\rangle + |O_p O_e\rangle)/\sqrt{3}$  (27). For the intermediate  $|I_p O_e\rangle$ , the most negative weak value will be given by a post-selected state which lies in the plane spanned by  $|\psi\rangle$  and  $|n\rangle$  as we noted in Sect. 3. We parametrize it as  $|\phi\rangle := \cos\theta|I_p O_e\rangle + \sin\theta\frac{|O_p I_e + O_p O_e\rangle}{2}$  by an angle  $\theta$ . It can be shown by a simple calculation that the weak value  $w(\phi) = \frac{\langle\phi|n\rangle\langle n|\psi\rangle}{\langle\phi|\psi\rangle} = \cos(\theta_n + \xi) \cos\theta_n / \cos\xi$  (19) can be  $-\infty$  for  $\xi = \pi/2$  and therefore  $\theta = -\pi/4$ , since  $\langle\phi|\psi\rangle = \cos\xi = (\sin\theta + \cos\theta)/\sqrt{3}$ .

On the other hand, the weak values for the orthogonal projection operators are

Weight	1/12	1/12	1/12	3/4	
$\phi$	$D_p D_e$	$D_p B_e$	$B_p D_e$	$B_p B_e$	Average
$I_p O_e$	1	-1	1	1/3	1/3
$O_p I_e$	1	1	-1	1/3	1/3
$O_p O_e$	-1	1	1	1/3	1/3
Sum	1	1	1	1	1

(50)

Let us check the consistency conditions 1 and 2 in the Table (50). The last row indicates that the consistency condition 1 (2) is satisfied, since the sum of the weak values over the complete set of orthonormal projectors is unity for all the post-selected states. To confirm the consistency 2 (8) we see in the rightmost column that the weighted average of the weak values over the post-selected states for each intermediate state becomes the probability for the intermediate state to be actually observed by projective measurements. Let us look at the column of  $D_p D_e$ . The apparent puzzle is that the “probabilities” for  $I_p O_e$  or  $O_p I_e$  add up to 2 exceeding 1. The weak value resolution to it is that the “probability” of the remaining  $O_p O_e$  is  $-1$  so that the total “probability” is  $1 + 1 - 1 = 1$ !

With this example we can see more clearly how the minus sign emerges in the table of the weak values. Suppose we have two non-orthogonal projection operators like  $I_p O_e + O_p I_e$  and  $O_p(I_e + O_e)$ . Assume that all the components of them are positive. Otherwise, we have done. Keep one of them as it is and find an orthogonal vector to it by the Schmidt orthogonalization of the other one. One may be easily convinced that the new one contains at least one negative component, unless the new one is lying along one of the axis. If the latter is the case we can go on picking up the third one and then

Schmidt diagonalization. If we end up with the configuration of the weak values which are all along the different axes, this is against our assumption of counter-factuality. That is, we could verify the weak value by a projective measurement. In Hardy's paradox, the counter-factuality is unavoidable so that the negative weak value should appear for a set of orthogonal set of projection operators and post-selected states.

Here we would like to re-emphasize the remark around Eq. (6) that the concept of weak value gives a physical basis of counter-factual arguments in Hardy's setting. The apparent paradox comes from the omission of the non-local possibility  $I_p O_e + O_p I_e$  and the non-orthogonality of  $O_p(I_e + O_e)$  and  $(I_p + O_p)O_e$ . Hopefully the weak value is helpful to resolve quantum paradoxes which come from counter-factual arguments in more general settings.

## 5. Spin Case

So far we have mainly studied the projection operators as observables, which are relevant in most of the optical experiments. In this section we study the non-commuting variables in the simple spin 1/2 case. In contrast to the preceding case, we will encounter complex weak values and a new consistency condition.

We choose the Pauli operators  $\sigma_x, \sigma_y$ , and  $\sigma_z$  as the three non-commuting observables. Let us start with an illustrative example: the pre-selected state is given by  $|\psi\rangle = |0_x\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$  and the post-selected states are  $\langle 0_y| = \frac{1}{\sqrt{2}}(\langle 0| + i\langle 1|)$ ,  $\langle 1_y| = \frac{1}{\sqrt{2}}(\langle 0| - i\langle 1|)$ . The table for the weak values is

Weight	1/2	1/2	
Post – selection	$\langle 0_y $	$\langle 1_y $	Average
$\sigma_x$	1	1	1
$\sigma_y$	1	-1	0
$\sigma_z$	$i$	$-i$	0

(51)

Table (51) exhibits the fact that the operators  $\sigma_x$  and  $\sigma_y$  are orthogonal with respect to the pre-selected state  $|\psi\rangle = |0_x\rangle$ , since  $\langle\psi|\sigma_x\sigma_y|\psi\rangle = i\langle 0_x|\sigma_z|0_x\rangle = 0$ . From the discussion of Sect. 2, we see that the weak vectors of  $\sigma_x$  and  $\sigma_y$  are real and orthogonal to each other, while that of  $\sigma_z$  is complex.

We can also compose the non-orthogonal positive weak value vectors, e.g., as

Weight	1/2	1/2	
Post – selection	$\langle 0_y $	$\langle 1_y $	Average
$\sigma_x$	1	1	1
$\sigma_x + \sigma_y$	2	0	1

(52)

We can clearly see that if the non-orthogonal vectors of weak values are positive vectors, then the orthogonal operators with respect to the preselected states contain at least one negative weak value.

We would like to see how strange weak values, which are outside the range of eigenvalues, show up in the spin 1/2 case for the more general pre-selected state

$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$  where  $\alpha$  and  $\beta$  are assumed to be real to simplify the exposition. The post-selected states are  $|0\rangle$  and  $|1\rangle$ . We compute the weak values,

Weight	$\alpha^2$	$\beta^2$	(Bloch vectors)
Post – selection	$\langle 0 $	$\langle 1 $	Average
$\sigma_x$	$\beta/\alpha$	$\alpha/\beta$	$2\alpha\beta$
$\sigma_y$	$-i\beta/\alpha$	$i\alpha/\beta$	0
$\sigma_z$	1	-1	$\alpha^2 - \beta^2$
Sum of squared	1	1	1

(53)

where “Sum of squared”  $:= ({}_f\langle\sigma_x\rangle_\psi^w)^2 + ({}_f\langle\sigma_y\rangle_\psi^w)^2 + ({}_f\langle\sigma_z\rangle_\psi^w)^2 = 1$ .

Of course, they are not orthogonal to each other. The weighted average gives the Bloch vector of the pre-selected state  $|\psi\rangle$  as shown in Table (53). The consistency with the multiplication rules,  $\sigma_x\sigma_y = i\sigma_z$  etc. can be checked, as well as the “Sum of squared” results.

Let us see the weak value for a pair of orthogonal projection operators,  $P_\pm := (1 \pm \sigma_x)/2$ :

Weight	$\alpha^2$	$\beta^2$
Post – selection	$\langle 0 $	$\langle 1 $
$P_+$	$\frac{1+\beta/\alpha}{2}$	$\frac{1+\alpha/\beta}{2}$
$P_-$	$\frac{1-\beta/\alpha}{2}$	$\frac{1-\alpha/\beta}{2}$
Sum	1	1

(54)

It is evident that at least one of the weak values is negative. For  $\beta/\alpha > 1$  the weak value of  $P_-$  with the post-selection  $\langle 0|$  is negative. This also means the weak value of  $\sigma_x$  for the post-selection  $\langle 0|$  is strange, i.e., more than unity and thus outside the range of the eigenvalue spectrum discussed in Sect. 3.

The appearance of strange values can also be seen without explicitly calculating the weak values. Note that  $({}_f\langle\sigma_x\rangle_\psi^w)^2 + ({}_f\langle\sigma_y\rangle_\psi^w)^2 + ({}_f\langle\sigma_z\rangle_\psi^w)^2 = 1$  in general for an arbitrary post-selected state  $\langle f|$ . This implies that  ${}_0\langle\sigma_x\rangle_\psi^w {}_1\langle\sigma_x\rangle_\psi^w + {}_0\langle\sigma_y\rangle_\psi^w {}_1\langle\sigma_y\rangle_\psi^w + {}_0\langle\sigma_z\rangle_\psi^w {}_1\langle\sigma_z\rangle_\psi^w = 1$ . By a suitable rotation, we will have  ${}_0\langle\sigma_x\rangle_\psi^w {}_1\langle\sigma_x\rangle_\psi^w = 1$ . We can see either  ${}_0\langle\sigma_x\rangle_\psi^w$  or  ${}_1\langle\sigma_x\rangle_\psi^w$  is strange, i.e., larger than 1. The ABL formula can be confirmed using  $|\langle 0|\psi\rangle|^2 |{}_f\langle 0_x\rangle\langle 0_x|\rangle_\psi^w|^2 = (\alpha + \beta)^2/4 = |\langle 0|0_x\rangle|^2 |\langle 0_x|\psi\rangle|^2$ . The intermediate states are orthogonal and therefore distinguishable so that they can be identified by a projective measurement, i.e., factual.

## 6. Summary

In this paper, we have developed a framework of weak value vectors and derived the two consistency conditions for them which enables us to interpret them as (complex) conditional probabilities for counter-factual processes in quantum paradoxes. We have shown that at least one of the components of the weak value vector becomes strange for counter-factual process. Two examples Hardy’s paradox and the spin 1/2 system were demonstrated. In the former, the choice of the pre-selected state is unique from



the setting of the paradox and the appearance of the “negative probability” is due to the counter-factuality. In the latter we have discussed the case of non-commutative operators to weakly observe.

We would like to emphasize that the weak value should be studied on its own right and believe that this is just the right quantity to describe a sequence of counter-factual phenomena in quantum mechanics by virtue of the “probabilistic” interpretation explained in this paper.

## Acknowledgments

The authors would like to acknowledge the useful comments of Professor Masanao Ozawa and the stimulating discussions with Professor Holger F. Hofmann. The authors are supported by Global Center of Excellence Program “Nanoscience and Quantum Physics” at Tokyo Institute of Technology. YS is also supported by JSPS Research Fellowships for Young Scientists (Grant No. 21008624).

## References

- [1] Aharonov Y, Albert D Z, and Vaidman L 1988 *Phys. Rev. Lett.* **60** 1351
- [2] Aharonov Y and Vaidman L 1991 *J. Phys. A: Math. Gen.* **24** 2315
- [3] Aharonov Y and Rohrlich D 2005 *Quantum Paradoxes* (Wiley-VCH, Weinheim)
- [4] Aharonov Y and Vaidman L 2008 in *Time in Quantum Mechanics*, Vol. 1, edited by Muga J G, Sala Mayato R, and Egusquiza I L (Springer, Berlin Heidelberg) p 399
- [5] Aharonov Y and Tollaksen J 2010 in *Visions of Discovery: New Light on Physics Cosmology and Consciousness*, edited by Chiao R Y, Cohen M L, Leggett A J, Phillips W D, and Harper Jr C L (Cambridge University Press, Cambridge, in press), arXiv:0706.1232
- [6] Ritchie N W M, Story J G, and Hulet R G 1991 *Phys. Rev. Lett.* **66** 1107
- [7] Parks A, Cullin D, and Stoudt D 1998 *Proc. Roy. Soc. Lond. A* **454** 2997
- [8] Resch K J, Lundeen J S, and Steinberg A M 2004 *Phys. Lett. A* **324** 125
- [9] Pryde G L, O’Brien J L, White A G, Ralph T C, and Wiseman H M 2005 *Phys. Rev. Lett.* **94** 220405
- [10] Hosten O and Kwiat P 2008 *Science* **319** 787
- [11] Resch K J 2008 *Science* **319** 733
- [12] Lundeen J S and Steinberg A M 2009 *Phys. Rev. Lett.* **102** 020404
- [13] Yokota K, Yamamoto T, Koashi M, and Imoto N 2009 *New J. Phys.* **11** 033011
- [14] Palacios-Laloy A, Mallet F, Nguyen F, Bertet P, Vion D, Esteve D, and Korotkov A N 2010 *Nature Phys.* **6** 442
- [15] Mitchison G, Jozsa R, and Popescu S 2007 *Phys. Rev. A* **76** 060302(R)
- [16] Jozsa R 2007 *Phys. Rev. A* **76** 044103
- [17] Reznik B and Aharonov Y 1995 *Phys. Rev. A* **52** 2358
- [18] Parks A 2003 *J. Phys. A: Math. Gen.* **36** 7185
- [19] Vaidman L 1996 *Found. Phys.* **26** 895
- [20] Aharonov Y and Botero A 2008 *Phys. Rev. A* **72** 052111
- [21] Dirac P A M 1942 *Proc. Roy. Soc. Lond. A* **180** 1
- [22] Feynman R P 1987 “Negative Probability” in *Quantum Implications : Essays in Honour of David Bohm* edited by Hiley B J and Peat F D (Routledge and Kegan Paul Ltd., London and New York) p 235
- [23] Hofmann H F 2010 *Phys. Rev. A* **81** 012103

- [24] Aharonov Y, Bergmann P G, and Lebowitz J L 1964 *Phys. Rev.* **134** B1410
- [25] Steinberg A M 1995 *Phys. Rev. Lett.* **74** 2405
- [26] Steinberg A M 1995 *Phys. Rev. A* **52** 32
- [27] Hardy L 1992 *Phys. Rev. Lett.* **68** 2981
- [28] Unruh W 1999 *Phys. Rev. A* **59** 126
- [29] Aharonov Y, Botero A, Popescu S, Reznik B, and Tollaksen J 2002 *Phys. Lett. A* **301** 130
- [30] Stapp H P 1997 *Am. J. Phys.* **65** 300
- [31] Mermin N D 1998 *Am. J. Phys.* **66** 920
- [32] Griffiths R B 1999 *Phys. Rev. A* **60** 5(R)
- [33] McCall S 2001 *Found. Phys. Lett.* **14** 95
- [34] Sokolovski D, Gimenez I P, and Sala Mayato R 2008 *Phys. Lett. A* **372** 3784
- [35] Tsang M 2010 *Phys. Rev. A* **81** 013824
- [36] Einstein A, Podolsky B, and Rosen N 1935 *Phys. Rev.* **47** 777
- [37] Johansen L M 2007 *Phys. Lett. A* **366** 374
- [38] Shikano Y and Hosoya A 2010 *J. Phys. A: Math. Gen.* **43** 025304
- [39] Ozawa M 2007 in *Proc. 8th Int. Conf. on Quantum Communication, Measurement and Computing* (NICT Press, Tokyo) p 363
- [40] Tollaksen J 2007 *Proc. SPIE* **6573** 65730H